Received: 7 February 2017 | Revised: 7 September 2017 | Accepted: 20 October 2017

DOI: 10.1002/gepi.22104

RESEARCH ARTICLE

WILEY Genetic Epidemiology

OFFICIAL JOURNAL INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY www.geneticepi.org

A robust and powerful two-step testing procedure for local ancestry adjusted allelic association analysis in admixed populations

Qing Duan[1,2,3] | Zheng Xu[1,4,5,7,8] | Laura M. Raffield[1] | Suhua Chang[9] | Di Wu[4,6] | Ethan M. Lange[10,11] | Alex P. Reiner[12,13] | Yun Li[1,2,4,5]

[1]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, United States of America

[2]Curriculum in Bioinformatics and Computational Biology, University of North Carolina, Chapel Hill, North Carolina, United States of America

[3]Department of Statistics, University of North Carolina, Chapel Hill, North Carolina, United States of America

[4]Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina, United States of America

[5]Department of Computer Science, University of North Carolina, Chapel Hill, North Carolina, United States of America

[6]Department of Periodontology, University of North Carolina, Chapel Hill, North Carolina, United States of America

[7]Department of Statistics, University of Nebraska-Lincoln, Lincoln, Nebraska, United States of America

[8]Initiative of Quantitative Life Sciences, University of Nebraska-Lincoln, Lincoln, Nebraska, United States of America

[9]CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[10]Department of Medicine, University of Colorado at Denver, Anschutz Medical Campus, Aurora, Colorado, United States of America

[11]Department of Biostatistics and Informatics, University of Colorado at Denver, Anschutz Medical Campus, Aurora, Colorado, United States of America

[12]Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America

[13]Department of Epidemiology, University of Washington, Seattle, Washington, United States of America

Correspondence
Yun Li, Department of Genetics, Campus Box 7264, University of North Carolina, Chapel Hill, North Carolina, 27599, United States of America.
Email: yunli@med.unc.edu

Funding information
National Human Genome Research Institute, Grant/Award Numbers: R01HG006292, R01HG006703; National Heart, Lung, and Blood Institute, Grant/Award Numbers: R01HL129132, R21HL126045

ABSTRACT

Genetic association studies in admixed populations allow us to gain deeper understanding of the genetic architecture of human diseases and traits. However, population stratification, complicated linkage disequilibrium (LD) patterns, and the complex interplay of allelic and ancestry effects on phenotypic traits pose challenges in such analyses. These issues may lead to detecting spurious associations and/or result in reduced statistical power. Fortunately, if handled appropriately, these same challenges provide unique opportunities for gene mapping. To address these challenges and to take these opportunities, we propose a robust and powerful two-step testing procedure Local Ancestry Adjusted Allelic (LAAA) association. In the first step, LAAA robustly captures associations due to allelic effect, ancestry effect, and interaction effect, allowing detection of effect heterogeneity across ancestral populations. In the second step, LAAA identifies the source of association, namely allelic, ancestry, or the combination. By jointly modeling allele, local ancestry, and ancestry-specific allelic effects, LAAA is highly powerful in capturing the presence of interaction between ancestry and allele effect. We evaluated the validity and statistical power of LAAA through simulations over a broad spectrum of scenarios. We further illustrated its usefulness by application to the Candidate Gene Association Resource (CARe) African American participants for association with hemoglobin levels. We were able to replicate

independent groups' previously identified loci that would have been missed in CARe without joint testing. Moreover, the loci, for which LAAA detected potential effect heterogeneity, were replicated among African Americans from the Women's Health Initiative study. LAAA is freely available at https://yunliweb.its.unc.edu/LAAA.

KEYWORDS

admixed populations, association analysis, effect heterogeneity, Genome-wide association studies, GWAS, local ancestry, population stratification, testing procedure

1 | INTRODUCTION

Genome-wide association studies (GWAS) have been successful in improving our understanding of the genetic basis of heritable diseases and quantitative traits (Visscher, et al., 2012). Although GWAS have initially been performed predominantly among individuals of European ancestry, the field has expanded to non-European and genetically admixed populations (Rosenberg, et al., 2010). Performing genetic association analysis in diverse populations allows us to gain deeper understanding of the genetic architecture of human diseases and traits, through assessing the generalizability of associated variants (Chen, et al., 2012; Ioannidis, Ntzani, & Trikalinos, 2004), narrowing down the location of the functional variants over the risk region (Helgason, et al., 2007; International HapMap, 2005) and identifying novel disease loci which are absent or in low frequency in European populations (Rosenberg, et al., 2010). Therefore, in the United States, genetically admixed populations have been receiving increasing attention. Despite these efforts, genetic association studies in admixed populations remain insufficient (Popejoy & Fullerton, 2016).

Admixed populations present unique challenges, namely complex LD patterns and the interplay of local ancestry (LA) and allele effect(s) on phenotypic traits. LD patterns are more complex in admixed populations resulting from a combination of ancestry-specific, historical recombination between variants in close proximity (genetic linkage) and the recent admixture process, where gene flow occurs among two or more distinct ancestral populations. Chromosomes of admixed individuals can be viewed as mosaics of segments originating from the contributing ancestral populations (Shriner, 2013). Therefore, admixed populations manifest two forms of LD–LD due to genetic linkage (ancestry LD) and that due to the admixture process (admixture LD) (Chakraborty & Weiss, 1988). The unique LD patterns enable admixture mapping, a method leveraging the extended admixture LD to scan for the association between phenotypic traits and chromosome segments of certain ancestry, with the assumption that the functional variants leading to increased risk have higher frequencies in chromosomal segments inherited from the ancestral population with higher disease prevalence (Chakraborty & Weiss, 1988; Shriner, 2013). Admixture mapping is particularly effective in detecting genetic loci with differential risk in the parental populations and requires the use of only a small panel of ancestry informative makers. Despite these advantages, the coarse scale of admixture LD results in low resolution for gene mapping (Winkler, Nelson, & Smith, 2010). With a dense collection of genotyped markers, conventional GWAS detects allelic association and provides higher resolution gene mapping. GWAS in admixed populations, however, pose special challenges due to the aforementioned presence of both ancestry LD and admixture LD, which may cause population stratification, thereby potentially leading to both spurious associations and false negatives (Kittles, et al., 2002; Liu, et al., 2013; Mao, et al., 2013; Qin, et al., 2010; Thornton & Bermejo, 2014; Wang, et al., 2011; Zhang & Stram, 2014). To address this challenge, it is necessary to properly adjust for both global and LA when conducting genetic association studies in admixed samples.

A number of advanced LA inference methods have been developed to infer LA down to the marker level from high throughput genotyping or sequencing data in admixed samples (Baran, et al., 2012; Patterson, et al., 2004; Price, et al., 2009; Wang, et al., 2014). Using the estimated LA for each marker and the derived global ancestry, association studies have been performed by adjusting for LA as a covariate (Levin, et al., 2014), by conducting Single Nucleotide Polymorphism (SNP) and admixture mapping separately using the same set of markers (Chen, et al., 2013; Reiner, et al., 2012) or by selecting a subset of African ancestry individuals for subsequent analysis (Li, et al., 2013).

The primary goal of all of these ancestry adjustment approaches is statistical validity or protected type-I error rate to reach the desired level of false positives. However, the other major challenge–the complex interplay of two sources of effects on phenotypic traits: LA and genotype allelic effects–is not explicitly considered. Therefore, the statistical power may not be optimal depending on the true known genetic architecture. Theoretically, allele and ancestry effects contain independent information, thus combining the two types of information may lead to increased statistical power. Tang et al. demonstrated the power gains in a joint testing procedure by integrating the two types of information in a

family-based study (Tang, et al., 2010). Later, Pasaniuc et al. developed a joint test (MIXSCORE) in the case-control GWAS setting (Pasaniuc, et al., 2011). This method has been applied to a fine-mapping study in African Americans (Levin, et al., 2014). A Bayesian joint test called BMIX was developed for admixture mapping and association mapping in unrelated individuals (Shriner, Adeyemo, & Rotimi, 2011). BMIX was shown to be generally more powerful than MIXSCORE. All these methods show the usefulness of combining allele and ancestry information assuming a consistent allele effect across the ancestral populations. However, effect heterogeneity may be present when analysis is conducted across diverse populations due to differential LD or differential haplotype effects among ancestral populations (Liu, et al., 2013). For example, with a sample size required to achieve > 80% power to detect significant associations, association analysis in African Americans fails to replicate the validated associations identified in Caucasians with the same allelic effect (Frazier-Wood, et al., 2013). To mitigate the problem of false negatives (missing genuine signals), Liu et al. explicitly included an interaction term between SNP allele effect and LA effect in a logistic regression model and showed increased power in the presence of substantial differential LD among ancestral populations (Liu, et al., 2013). Recently, McHugh et al. developed a method called CAnD to detect heterogeneity in population structure across the genome in admixed populations (McHugh, Brown, & Thornton, 2016).

In this study, we propose a robust and powerful two-step testing procedure for association studies in admixed populations. In the first step of the testing procedure, we jointly test allele effect, ancestry effect, and allele effect heterogeneity across ancestry in a regression model. The omnibus joint test guards against missing genuine associations from any of the three sources. This feature is important because the true underlying genetic architecture at each locus is unknown. Similar model selection problems are common in genetic association studies (Conneely & Boehnke, 2007). Loci manifesting significance in the first step are carried on to the second step, where we determine the source of association through a one-time model selection process. We model the interaction between allele and ancestry through an ancestry-specific allele effect. This approach not only captures the effect heterogeneity among ancestral populations, but is also more powerful than simply modeling the interaction using a cross-product term, particularly when LA is estimated. In the present study, we assess the power and type I error of the proposed testing procedure together with existing approaches by conducting extensive simulations mimicking a broad spectrum of real-life scenarios. The scenarios considered range from one extreme of solely an allelic effect, to the other extreme of solely an ancestry effect, to anywhere in between, as well as differential allelic effects across LA groups. In addition to using *true* simulated LA, we demonstrate the robustness of our results using *estimated* LA in simulated data, which is more realistic since LA is not directly observed in real data, but rather inferred statistically. Finally, we apply our two-step testing procedure to perform genome-wide association analysis for hemoglobin levels among African American participants in the Candidate Gene Association Resource (CARe) (Musunuru, et al., 2010), where we identified effect heterogeneity at a locus previous known to be associated with hemoglobin levels. We replicated the previous finding (Lo, et al., 2011) that the minor allele shared between individuals of European and those of African ancestry is associated with lower hemoglobin levels. More importantly, we identified effect heterogeneity that the minor allele from African ancestry was associated with even lower hemoglobin levels. We replicated this effect heterogeneity in an independent dataset from the Women's Health Initiative (WHI) Study.

2 | METHODS

2.1 | Simulation of admixed samples and reference haplotypes

We simulated diploid samples of individuals with admixed African and European ancestry, as well as African and European reference haplotypes using COSI (Schaffner, et al., 2005), a sequence data simulator based on coalescent model. COSI is well calibrated to generate genetic data that closely resemble empirical data in allele frequency, LD and population differentiation (Schaffner, et al., 2005). We first generated 3,000 African haplotypes and 3,000 European haplotypes, 50 kb each, to serve as the ancestral haplotypes. Then 1,000 haplotypes from each of the ancestral populations were randomly selected and kept as reference haplotypes. The remaining 2,000 African and 2,000 European haplotypes were randomly combined to generate 1,000 African American samples. Sampling without replacement was adopted such that none of the parental haplotypes were reused. We constructed admixed samples based on the empirical estimate of 0.2125 switch points per 50 kb region (Wegmann, et al., 2011). To be specific, we first generated 425 African American mosaic chromosomes each containing one switch point by joining one randomly selected African chromosomal segment and one randomly selected European chromosomal segment. The switch point assignment was based on the recombination rate from the crossover map generated by COSI. We used a binomial distribution with probability 0.5 to decide which ancestral population was the first/left segment. Next, from the pool of unused chromosomes, we randomly selected 1,408 African and 167 European, together with the 425 mosaic chromosomes generated above, to generate a collection of 2,000 chromosomes. The rates of switch point occurrence and proportion (80%) of African ancestry chromosomes were consistent with findings from previous work (Parra, et al., 1998;

TABLE 1 Four simulation settings

Scenario	Ancestry effect (γ)	Allelic effect (β)	Effect direction (AFR/EUR)	Allele frequency difference (δ)
LA[1]	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1	0	NA/NA	0–0.05, 0.05–0.1, 0.1–0.2, 0.2–0.3, 0.3–0.4, and > 0.4
AA[2]	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1	$\gamma \times k$[5]	+ / NA	0–0.05, 0.05–0.1, 0.1–0.2, 0.2–0.3, 0.3–0.4, and > 0.4
GA[3]	0	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1	+ / +	0–0.05, 0.05–0.1, 0.1–0.2, 0.2–0.3, 0.3–0.4, and > 0.4
EH[4]	0	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1	+ / -	0–0.05, 0.05–0.1, 0.1–0.2, 0.2–0.3, 0.3–0.4, and > 0.4

[1]LA: Local ancestry only effect); [2]AA: Local ancestry and allelic effect both exist; [3]GA: Genotype allelic only effect; [4]EH: Effect heterogeneity; [5]$k = 0.1$, when $0 < \delta < 0.05$; $k = 0.2$, when $0.05 < \delta < 0.1$; $k = 0.3$, when $0.1 < \delta < 0.2$; $k = 0.4$, when $0.2 < \delta < 0.3$; $k = 0.5$, when $0.3 < \delta < 0.4$; and $k = 0.6$, when $\delta > 0.4$.

Wegmann, et al., 2011). Finally, we randomly paired the 2,000 haploid chromosomes to yield 1,000 diploid African American samples. We repeated the process 1,000 times to create 1,000 regions.

2.2 | Association tests

We evaluated the performance of the first step of our proposed testing procedure (LAAA: Local Ancestry Adjusted Allelic) and three other existing methods genotype-allelic-only (GAO) (Li, et al., 2013), local-ancestry-only (LAO) (Chen, et al., 2013), and Ancestry Plus Allelic (APA) (Levin, et al., 2014)) by comparing power and type I error in the four simulated scenarios detailed below and tabulated in Table 1. Under scenario EH, where effect heterogeneity is present, we compared our LAAA model with a Bayesian joint test (BMIX) (Shriner et al., 2011) and the conventional interaction model with a cross-product interaction term conventional interaction model (CPM) (Liu, et al., 2013).

- GAO: $E(Y) = \alpha_0 + \alpha_{E_1} E_1 + \alpha_{E_2} E_2 + \alpha_G G + \beta X_{ref}$, where E_1 and E_2 are covariates; G is the estimated global African ancestry (half of the average copies of African allele per person); and X_{ref} is the number of reference alleles at the locus under investigation. We test the null hypothesis of $\beta = 0$. This GAO test is the one commonly used in GWAS to examine whether there is an allele effect on the phenotype, assuming an additive allelic model.
- LAO: $E(Y) = \alpha_0 + \alpha_{E_1} E_1 + \alpha_{E_2} E_2 + \alpha_G G + \gamma X^{afr}$, where X^{afr} is the number of African ancestry alleles at the locus under investigation, and all other notations are the same as in the GAO model above. We test the null hypothesis of $\gamma = 0$. LAO is the statistical test used in admixture mapping to scan for an ancestry effect of variants with frequency disparity across ancestry populations.
- APA: $E(Y) = \alpha_0 + \alpha_{E_1} E_1 + \alpha_{E_2} E_2 + \alpha_G G + \beta X_{ref} + \gamma X^{afr}$. We test the null hypothesis of $\beta = \gamma = 0$. APA jointly tests for allele and ancestry effects in an additive manner.
- Local Ancestry Adjusted Allelic (LAAA): $E(Y) = \alpha_0 + \alpha_{E_1} E_1 + \alpha_{E_2} E_2 + \alpha_G G + \beta X_{ref} + \gamma X^{afr} + \eta X^{afr}_{ref}$, where X^{afr}_{ref} is the number of African-ancestry reference alleles at the locus under investigation. LAAA is an extension of the APA model above by modeling the interaction between allelic copy and LA through a LA-specific allele term ηX^{afr}_{ref}. Utilizing the joint distribution of allele and LA provided by HAPMIX, we parameterize the interaction as the ancestry-specific allele estimate and propose a two-step testing procedure. In the first step, we test the null hypothesis of $\beta = \gamma = \eta = 0$. Since the true underlying causal variants and the LD between marker under study and the causal variant(s) are both unknown, testing all three terms simultaneously is the most robust approach, with minimal power loss, when the model is misspecified.
- BMIX: As a Bayesian joint test of ancestry and association, BMIX computes the posterior probability that a locus is associated with a trait by converting the P-value for the pooled estimates of genotype effects combined over strata of LA. The prior distribution of genotype effects is specified using the posterior probability from admixture mapping. Details can be found in (Shriner et al., 2011).
- CPM: $E(Y) = \alpha_0 + \alpha_{E_1} E_1 + \alpha_{E_2} E_2 + \alpha_G G + \beta X_{ref} + \gamma X^{afr} + \eta X_{ref} * X^{afr}$. A conventional cross-product interaction term between SNP allele effect and LA is included in the model, as described by Liu et al. (Liu, et al., 2013), to capture the presence of effect heterogeneity.

Once an SNP is identified to be significantly associated with the trait from the first step of LAAA, the second step of LAAA can be applied to identify the source of the association, i.e., to examine whether the association is driven primarily by SNP allele effect, ancestry effect, or the presence of effect heterogeneity. In the second step of LAAA, we perform a one-time model selection based on the test statistics associated with the SNP allele effect (β), ancestry effect (γ), and the ancestry-specific allele effect (η). Specifically, we identify the primary driving source of association by

selecting the effect with the largest absolute value of the test statistics.

2.3 | Simulation of quantitative phenotypic traits

For each locus, we simulated quantitative traits for the 1,000 African American samples based on a null model or a causal model. Our null model consists of two independent covariates, $E1$ and $E2$. Let $QT_i = 0.5E_{1i} + 0.5E_{2i} + \epsilon_i$, where $E_1 \sim Bernoulli(0.5)$ and $E_2 \sim Normal(0, 1)$. Error term follows a standard normal distribution and is independent across individuals. The causal models are detailed in the previous section.

As summarized in Table 1, we simulated four scenarios for the underlying causal model. In scenario LAO (local ancestry only), only a LA effect is present, with corresponding effect size, γ, ranging from 0.1 to 1 and no allele effect (i.e., $\beta = 0$). In Scenario AA (local ancestry, LA and allele effect), both LA and genotype allelic effects exist. Specifically, an allelic effect is driven by an ancestry effect with a factor of k. The value of k depends on the minor allele frequency (MAF) difference (δ) for the studied variant between African and European derived chromosomes. Scenario GA (genotype allelic only) has only a genotype allelic effect with effect size, β, ranging from 0.1 to 1 and no ancestry effect (i.e., $\gamma = 0$). In addition, we model the allelic effect being in the same direction across the two ancestral populations. Scenario effect heterogeneity (EH) was generated under CPM where allelic effect from African ancestry is in the opposite direction of that from European ancestry (γ equals 0 and β ranges from 0.1 to 1).

We evaluated the power of our proposed LAAA and alternatives tabulated above using both true/simulated and estimated LA. To be specific, in each of the 1,000 regions, 10 markers were randomly selected with replacement for each allele frequency difference category (six allele frequency difference categories in total, see Table 1). Then causal models were applied to generate quantitative traits under the alternative hypothesis. Power was calculated by counting the number of times in the 10,000 replications when P-value is less than the GWAS significance threshold of 5×10^{-8} for GAO, APA, LAAA, and CPM; for LAO model, we used the conventional significance threshold of 7×10^{-6} for admixture mapping (Pe'er, et al., 2008). For BMIX, since significance threshold for admixture mapping was calculated by $\frac{0.05}{\text{admixutre burden} \times 2}$ and association mapping by $\frac{0.05}{\text{association burden} \times 2}$, we, correspondingly, set the testing burden of admixture mapping to be 3571.429 ($\frac{0.05}{7 \times 10^{-6} \times 2}$) and the testing burden of association mapping to be 5×10^5 ($\frac{0.05}{5 \times 10^{-8} \times 2}$). Type I error was assessed based on 500,000 experiments. Specifically, in each of the 1,000 regions, 500 markers were randomly selected with replacement for each allele frequency difference category. Then the null model was used to generate the quantitative traits under the null.

To evaluated the performance of the second step of LAAA, that is, the proportion of times when the source of association can be correctly identified, we traced the source of association through a one-time model selection by comparing the absolute value of the test statistics associated with β, γ, and η among the significant loci from the first step.

2.4 | Inference of LA using HAPMIX

We used HAPMIX to infer LA with reference from the 1000 Genomes Project (Phase I, March 2012 release). HAPMIX provides highly accurate LA estimates in admixed samples by leveraging within-population LD based on a hidden Markov model (HMM) (Price, et al., 2009). We used the default parameters and "Diploid" mode. Instead of obtaining, by default, the expected probability of 0, 1, or 2 copies of European ancestry at each SNP, we obtained the inferred joint distribution of LA and genotype by setting "output details" to "prob" (see Figure S1 for an example). The probabilities from the joint modeling allowed us to calculate the expected copies of the reference allele (i.e., estimated genotype dosage, ranging continuously from 0 to 2), expected copies of African ancestry allele (i.e., estimated local African ancestry, ranging continuously from 0 to 2), and expected copies of African ancestry reference allele (again ranging continuously from 0 to 2). Since the 16 probabilities (reflecting the number of possible combinations of ordered genotypes (4) by ordered allele-ancestry combinations (4) for the two inherited alleles) of each marker may not sum up to 1, we normalized the probabilities, imposing the sum-up-to-one constraint. Accuracy of HAPMIX LA estimation was evaluated by calculating the Pearson correlation between the estimated and true LA.

2.5 | CARe data set

The CARe consortium has been previously described (Musunuru, et al., 2010). We applied LAAA to study associations in 5,711 African American participants from two CARe cohorts, namely Atherosclerosis Risk in Communities and Coronary Artery Risk Development In young Adults. These cohorts have been previously described (Bild, et al., 2002; Friedman, et al., 1988). All samples were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 Chip at the Broad Institute of MIT and Harvard. Markers with genotype call rate < 90%, Hardy–Weinberg exact test P-value $< 1 \times 10^{-6}$, or MAF < 1% were removed.

2.6 | WHI SNP Health Association Resource (WHI-SHARe) data set

We replicated our findings in a cohort of 8,087 African American participants from WHI-SHARe study. All samples

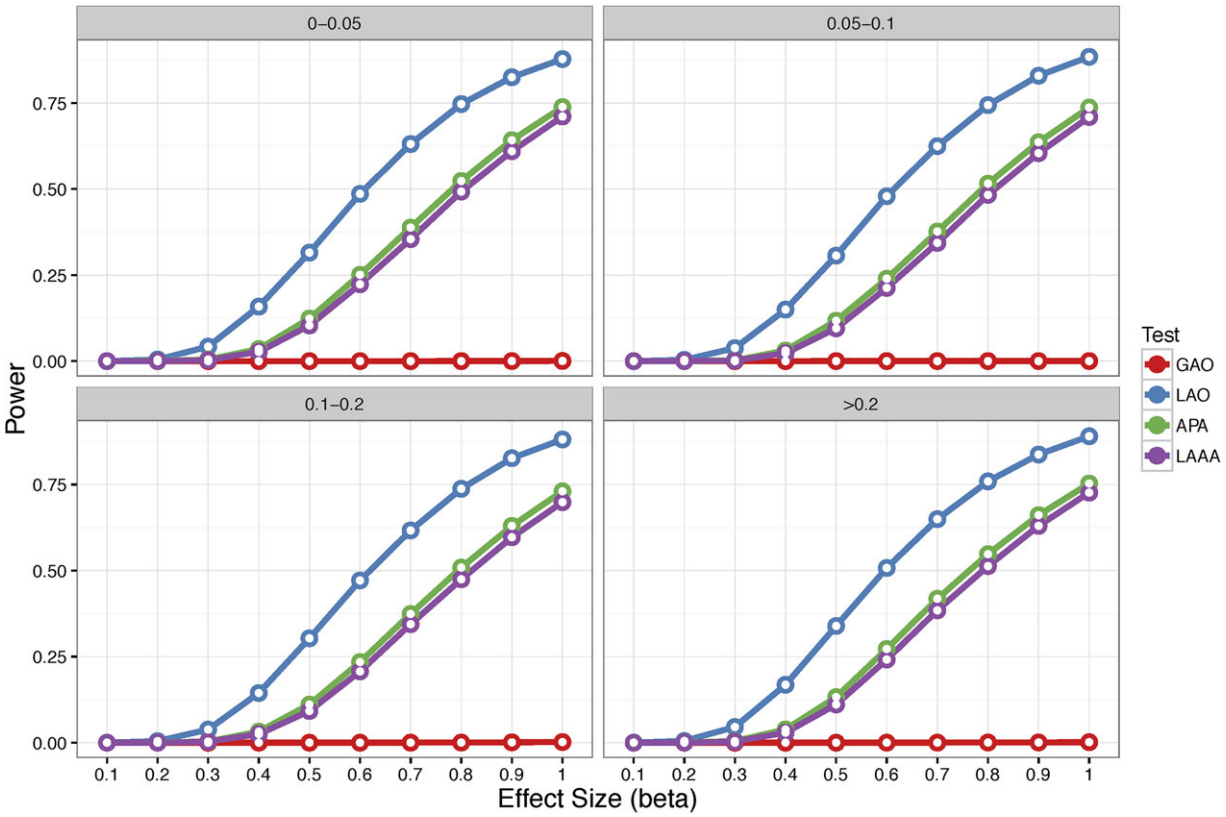

FIGURE 1 Statistical power of the four tests in scenario LA (LAO effect) where African ancestral alleles associated with trait. Power is plotted as a function of LA effect size stratified by allele frequency differences between African and European ancestral populations (GAO: Genotype-allelic-only; LAO: Local-ancestry only; APA: Ancestry plus allelic; LAAA: Local ancestry adjusted allelic)

were genotyped using the Affymetrix 6.0 genotyping platform. Prior to LA inference, we removed Affymetrix 6.0 SNPs with genotype call rate < 90%, or Hardy–Weinberg exact test *P*-value 1 $\times 10^{-6}$, or MAF < 1%. Quality control details were described previously (Reiner, et al., 2012; Reiner, et al., 2011).

3 | RESULTS

3.1 | Power evaluation with simulated data using true/simulated LA

An advantage of performing association analyses in admixed populations is that it allows the identification of risk variants leading to disease disparities, particularly variants with substantial allele frequency differences (a risk variant in one population is sometimes monomorphic in other populations) or risk variants with different risk alleles across populations. Examples of such variants include the *DARC* null allele associated with white blood cells (Nalls, et al., 2008), *SLC24A5* variant *Ala111Thr* associated with skin pigmentation (Lamason, et al., 2005) and the *APOL1*G1 and G2 variants in African-descent populations associated with kidney disease (Genovese, et al., 2010). Such associations may be missed if genetic analyses are conducted solely in one homogeneous population.

Bearing this in mind, we simulated Scenario LA, where the associations are present in only one parental population. In this scenario, we assumed that the causal risk allele, leading to an elevated mean trait value, is specific to African ancestry with a fixed allele frequency of 1 in African descendent populations and a frequency of 0 in European descent populations. We also assumed that the SNP under study is not causal, but is located near and in complete ancestry LD with the causal variant (the allele frequency differences between the two populations at the two variants result in admixture LD). Under this scenario, the SNP-trait association is driven entirely by the LA of the tested SNP (i.e., within African and European homogeneous populations the tested variant would not associated with the trait because the causal variant is monomorphic). As expected, LAO, APA, and LAAA, which are all designed to test for an ancestry effect, are all well powered to detect this association with modest effect size, regardless of the allele frequency difference between the two contributing populations (Figure 1). LAO has an advantage in this setting compared to APA and LAAA mainly due to the lower multiple-testing burden. In this scenario, statistical power is not affected much by allele frequency differences between the two populations, as the association is driven by LA rather than

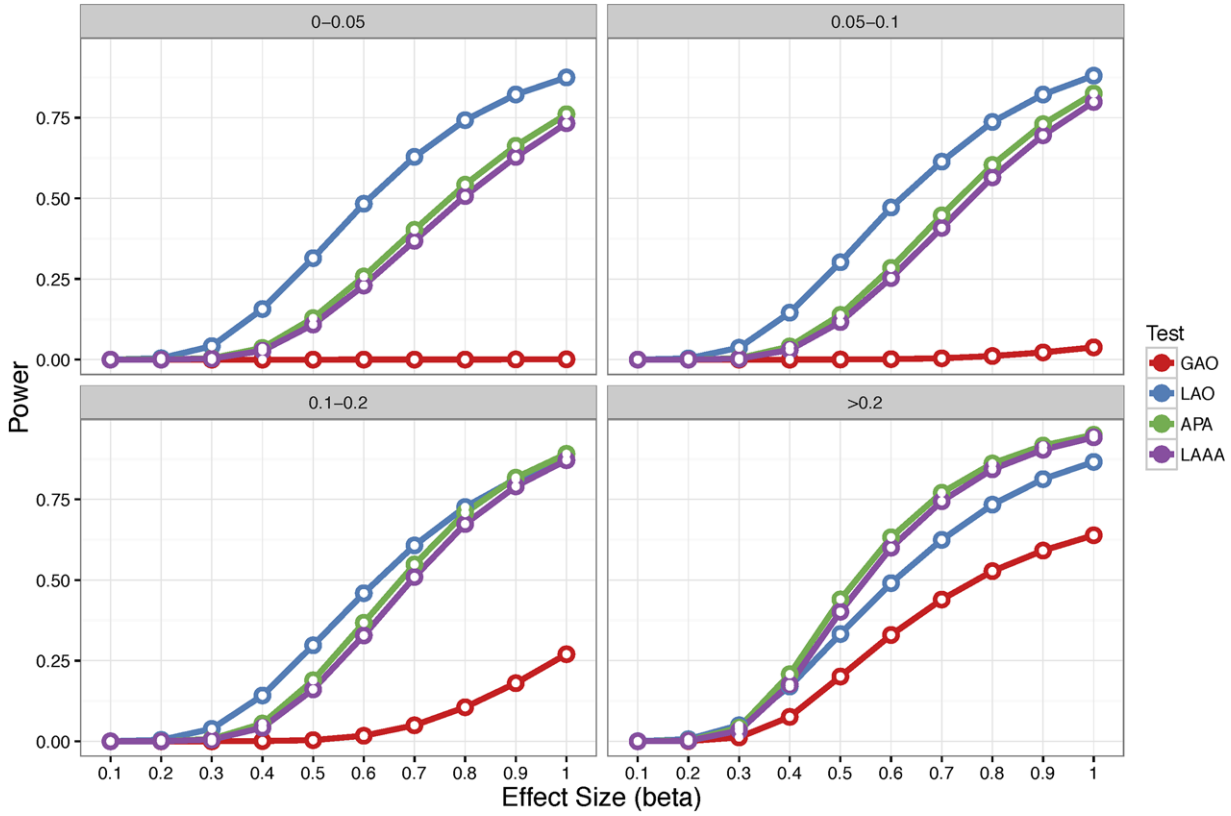

FIGURE 2 Statistical power of the four models in scenario AA (local ancestry and allelic effects both exist) where risk alleles present only in the African ancestral population with weak to moderate ancestry LD with the tested SNP. Power is plotted as a function of effect size stratified by allele frequency difference between African and European ancestral populations (GAO: Genotype-allelic-only; LAO: Local-ancestry-only; APA: Ancestry plus allelic; LAAA: Local Ancestry Adjusted Allelic)

allelic status. For example, with allelic effect $\beta = 0.7$, power is 0.01%, 62%, 37%, and 34% for the four tests GAO, LAO, APA, and LAAA, respectively when allele frequency difference is between 0.1 and 0.2 and 0.014%, 65%, 42%, and 38%, when allele frequency difference is > 0.2.

In addition to scenario LA, we simulated a more realistic case (Scenario AA) to allow for an allelic effect. Similar to the LA scenario, the causal allele is present only in the African ancestral population with a fixed allele frequency of 1, contributing to elevated mean trait values in this population. Again, the tested noncausal SNP is in strong admixture LD with the causal variant. In scenario AA, we assumed that the tested SNP is in low to moderate ancestry LD with the causal allele. The strength of the ancestry LD is reflected by allele frequency differences of the tested SNP in African and European populations. Specifically, a similar minor allele frequency in the two parental populations indicates low LD between the causal and the tested SNP, a case similar to that in scenario LA. As the frequency difference increases, the ancestry LD between the causal and the tested SNP becomes stronger, and we should observe both an ancestry and allelic effect. Thus, the power to capture the allelic effect depends on the allele frequency difference. As shown in Figure 2, when the allele frequency difference is small (0–0.03), the observed pattern is similar to that in scenario LA. The power of LAO does not change with the increase of the allele frequency difference, as it only tests for a LA effect. APA and LAAA, which test for both an allelic and ancestry effect, have increased power with increasing allele frequency differences and become more powerful than the LAO test for moderate effect size when the allele frequency difference is greater than 0.2. APA and LAAA have comparable power in this AA scenario.

Variants identified in European populations are often transferable to other ancestries with consistent effect direction and magnitude (Loth, et al., 2014; Teslovich, et al., 2010). Therefore, in scenario GA, we simulated a case where African and European ancestral populations share the same risk allele with the same direction and similar effect size. We anticipate such a GA scenario where LD between the causal and the tested SNP is also in the same direction across the ancestral populations. Difference in mean trait values across populations is caused by the difference in allele frequencies at the causal SNP. Figure 3 shows the power comparison under scenario GA for allelic effect size ranging from 0.1 to 1 stratified by allele frequency difference. GAO, APA, and LAAA have comparable performance over the entire range of effect sizes. The 1-*df* GAO test is slightly more powerful than the 2- and 3-*df* test (APA

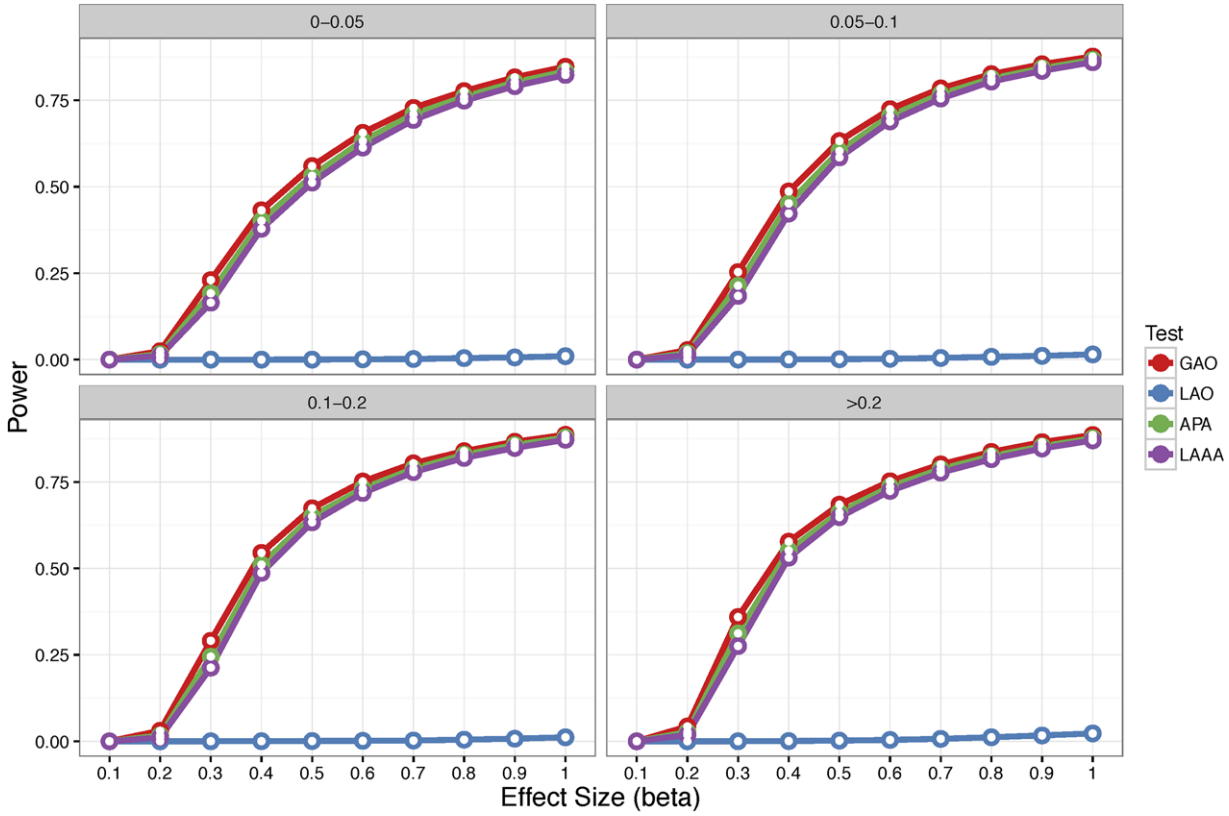

FIGURE 3 Statistical power of the four models in scenario GA (genotype allelic only effect) where risk alleles associate with trait both in African and European ancestral population with consistent effect size and direction. Power is plotted as a function of effect size stratified by allele frequency differences between African and European ancestral population (GAO: Genotype-allelic-only; LAO: Local-ancestry only; APA: Ancestry plus allelic; LAAA: Local Ancestry Adjusted Allelic)

and LAAA). As expected, LAO incurs dramatic power loss, as it fails to capture the allelic effect. For a fixed effect size, power tends to increase with the increasing differential allele frequencies.

For some variants, effect heterogeneity has been reported across populations, which motivated us to simluate scenario EH (Frazier-Wood, et al., 2013). In this scenario, a genuine association is present in each ancestral population but the LD between the causal and the tested SNP across ancestral populations is different or even in opposite directions, leading to effect heterogeneity of the tested SNP across populations. Although this phenomenon may affect only a small fraction of loci, it warrants consideration in order to obtain a more complete understanding of the underlying genetic structure in worldwide populations. We compared the power across a broad spectrum of effect sizes and allele frequency differences (Figure 4). In this scenario, LAAA outperforms all other tests. For instance, with a moderate effect size ($\beta = 0.5$) and for markers with allele frequency difference between 0.05 and 0.1, the power of LAAA is 0.60, corresponding to a relative power gain of 168%, 589%, and 150% when compared with GAO, LAO, and APA, respectively.

Once a significant signal is found by LAAA's omnibus test in the first step, it is important to identify the primary component contributing to the association. Thus, in the second step of LAAA, we aim to narrow down the primary source of association by performing a single model selection through comparing the absolute values of the test statistics of the allelic, ancestry, and ancestry-specific allelic effects. To examine the power of this second step, we calculated the fraction of times that an effect is correctly identified in each of the four scenarios, conditional on the tested SNP passing the significance threshold in the first step. Table 2 shows the average proportion of times that the source of association is correctly identified across small ($\beta = 0.3$) to large ($\beta = 1$) effect size. It is noteworthy that when effect heterogeneity exists (scenario EH), this step is powerful to identify the ancestry-specific risk allele with > 90% power (mean power 93.3%). In addition, it is efficient (power > 95%) for identifying ancestry effect as the only source of association (scenario LA) and ancestry with allele effect (scenario AA). The power to detect allelic effect as the only source of association (scenario GA) is slightly lower, but still above 80%. Table S1 shows the proportions broken down by effect size.

Importantly, modeling interaction of allelic status and LA by using the joint distribution of ancestry and allele (LAAA) is highly powerful when effect heterogeneity presents. Because, in LAAA, effect of heterogeneity across ancestral

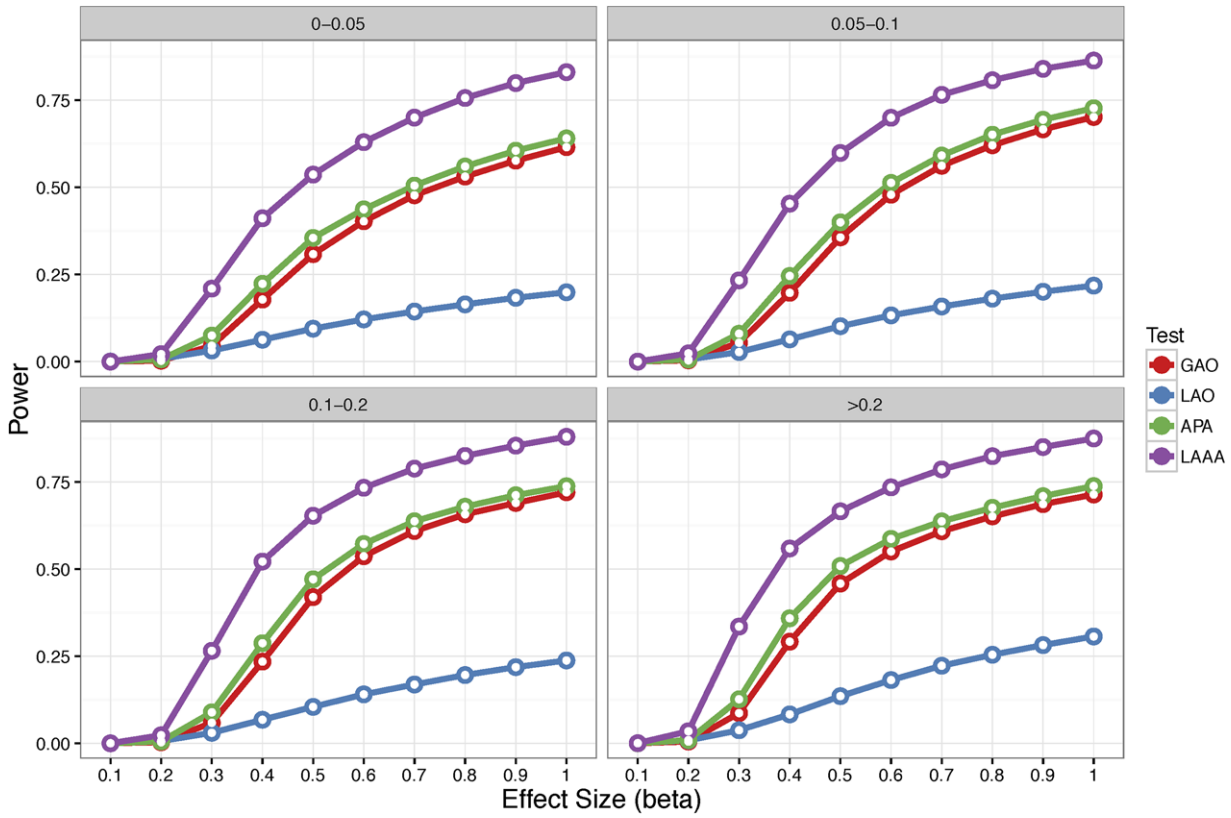

FIGURE 4 Statistical power of the four models in scenario EH (effect heterogeneity) where risk alleles associate with trait both in African and European ancestral populations with inconsistent effect direction. Power is plotted as a function of effect size stratified by allele frequency difference between African and European ancestral populations (GAO: Genotype-allelic-only; LAO: Local-ancestry only; APA: Ancestry plus allelic; LAAA: Local Ancestry Adjusted Allelic)

TABLE 2 Average proportion of times that the source of association is correctly identified

	Allele frequency difference			
	0–0.05	**0.05–0.1**	**0.1–0.2**	**>0.2**
Scenario LA	98.7%	99.6%	99.6%	99.1%
Scenario AA	99.5%	99.3%	98.1%	90.8%
Scenario GA	88.1%	87.8%	87.9%	89.3%
Scenario EH	93.5%	93.0%	93.2%	94.0%

LA: Local ancestry only effect; AA: Local ancestry and allelic effect both exist; GA: Genotype allelic only effect; EH: Effect heterogeneity.

TABLE 3 Ratio[1] of the statistical power of the joint test in LAAA to that of BMIX under Scenario EH

	Allele frequency difference			
B	**0–0.05**	**0.05–0.1**	**0.1–0.2**	**>0.2**
0.1	NA	NA	NA	2.00
0.2	1.99	2.28	1.99	1.94
0.3	2.25	2.35	2.39	2.11
0.4	1.76	1.78	1.75	1.50
0.5	1.62	1.61	1.47	1.4
0.6	1.75	1.60	1.47	1.55
0.7	1.89	1.61	1.63	1.76
0.8	1.96	1.64	1.83	1.90
0.9	1.96	1.72	1.94	1.96
1.0	1.96	1.83	2.05	1.97

[1]The ratio was calculated by dividing the power of LAAA by that of CPM. Values greater than 1 indicate LAAA having higher power than CPM. NA values indicate power is zero in both methods.

populations is captured by the number of ancestral specific reference alleles. The additional information of ancestry specific allelic count leads to enhanced power in the presence of effect heterogeneity. We performed power comparisons of our joint test (the first step of LAAA) with the Bayesian joint test BMIX (Shriner et al., 2011) and the conventional interaction model (CPM) (Liu, et al., 2013). Tables 3 and 4 display the ratios of the power of LAAA under scenario EH to that of BMIX and CPM, respectively. The ratios were calculated by dividing the power of LAAA by that of BMIX or CPM such that values > 1 indicate LAAA having higher power than BMIX or CPM. As shown in Tables 3 and 4, across different β values and allele frequency differences, we observe power advantages of our LAAA method: ratios are greater than 1. These results suggest that the joint test in LAAA, which accounts for the joint distribution of allele and ancestry effect, is more powerful than BMIX and CPM in capturing

TABLE 4 Ratio[1] of the statistical power of LAAA to that of CPM under Scenario EH

	Allele frequency difference			
B	**0–0.05**	**0.05–0.1**	**0.1–0.2**	**>0.2**
0.1	NA	NA	NA	10.29
0.2	2.53	2.63	2.82	2.46
0.3	1.75	1.80	1.89	1.69
0.4	1.31	1.30	1.27	1.19
0.5	1.19	1.16	1.12	1.11
0.6	1.14	1.11	1.09	1.09
0.7	1.11	1.08	1.07	1.07
0.8	1.08	1.06	1.05	1.06
0.9	1.07	1.05	1.04	1.04
1.0	1.06	1.04	1.04	1.04

[1]The ratio was calculated by dividing the power of LAAA by that of CPM. Values greater than 1 indicate LAAA having higher power than CPM. NA values indicate power is zero in both methods.

differential SNP allele effects. Additionally, we observed that the joint test in LAAA is more powerful than CPM particularly when effect sizes are small (Table 4).

3.2 | Type I error

We used 500,000 replicates under the null hypothesis to evaluate the type I error at the nominal significance level $\alpha = 0.05$, 0.01, and 0.001. Results are summarized in Table 5 and Figure 6. Across allele frequency difference categories, LAAA appropriately controlled the type I errors.

3.3 | Power in simulations with estimated LA

In the above fourscenarios, particularly Scenario EH, we show, using simulated data, that our LAAA testing procedure incorporating ancestry-specific allele counts is powerful when using true LA information. In real data analysis, however, LA and ancestry-specific allele counts are both unknown. Thus, the usefulness of the proposed tests is affected by the accuracy of the estimated LA, and the estimated joint distribution of allelic status and LA. For this reason, we examined the extent of power loss when using inferred ancestry.

We first evaluated the accuracy of the estimated LA and ancestry-specific allele designations by calculating the Pearson correlation coefficient between the true and inferred LA. As shown in Table 6, the median correlation between true and estimated genotype, African allele andAfrican reference allele are all above 0.8. Note that the inferred African reference allele is highly accurate. For markers with MAF less than 0.05 in the ancestral African population, we notice a decrease in the accuracy of ancestry-specific allele estimates but the median correlation is still above 0.8.

Having confirmed that LA can be inferred with reasonable accuracy at both the marker and allele level, we performed power analyses using inferred LA with the same simulated data set. Then, we compared the results with those obtained earlier using true ancestry. Overall, the patterns remain similar to the prior observations when using inferred LA in scenario LA through EH. Figure 5 shows the power evaluation for scenario EH. As expected, although using inferred ancestry incurs slight reductions in power, the patterns observed using true LA remain.

Compared to CPM, LAAA achieves higher power using estimated LA (Table S4). Note that the better performance of LAAA is more evident when using estimated LA than true LA. This may be due to the more accurate estimation of ancestry-specific reference allele than total ancestry-specific allele as shown in Table 6. We also examined the fraction of times the source of association is correctly identified using inferred LA and ancestry-specific allele in the second step of our LAAA. Similar results are obtained as those using true ancestry (Table S2 and Table S3).

3.4 | Application to real phenotypes

We applied LAAA to test for the association between hemoglobin level and genetic variants in African American samples from the CARe project. Many genetic loci have been reported to be associated with hemoglobin in Europeans (van der Harst, et al., 2012), East Asians (Kamatani, et al., 2010) and African Americans (Auer and Teumer, 2014; Auer, et al., 2012; Chen, et al., 2013; Lo, et al., 2011). We first performed a genome-wide 3-*df* joint test (LAAA first step) adjusting for age, estimated proportion of global African ancestry, cohort, and smoking status. Then, in the second step, we conducted a single model selection procedure comparing the test statistics of the allelic, ancestry, and ancestry-specific allele effects. We used a suggestive *P*-value threshold of 1×10^{-6}. Our analysis strategy replicated previously reported hemoglobin-associated loci. As shown in Table 7, SNPs rs9940149 and rs7199221 on chromosome 16 show a strong allelic effect, indicating similar effects in African and European ancestral populations. It is worth noting that rs1211375 on chromosome 16 shows a strong ancestry-specific allelic effect, suggesting effect heterogeneity at this locus in African and European ancestral populations. Rs1211375, in the alpha-globin (HBA2-HBA1) locus, is reported to be a common variant in both Europeans and African populations. Consistent with the results observed by Lo et al. (2011), the minor allele, which is shared between individuals of European and African ancestry, is associated with lower hemoglobin levels. However, we identified effect heterogeneity that the minor allele inherited on African-derived chromosomes was associated with even lower hemoglobin levels than when inherited on European-derived chromosomes (Table 7). We replicated this effect heterogeneity in an independent dataset from the WHI Study (*P*-value = 5.78×10^{-6}, allelic effect = 0.54,

TABLE 5 Ratio of type I error to the nominal significance level of 0.05, 0.01 and 0.001

Allele frequency difference	0.05				0.01				0.001			
	GAO	LAO	APA	LAAA	GAO	LAO	APA	LAAA	GAO	LAO	APA	LAAA
0–0.05	1.00	1.00	1.00	1.00	1.01	1.01	1.00	1.00	1.08	0.95	0.98	0.95
0.05–0.1	0.99	1.01	1.00	1.00	0.99	1.02	1.00	1.00	0.97	0.99	1.04	0.94
0.1–0.2	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99	0.95	0.91	0.92	0.92
>0.2	1.00	1.00	1.00	1.00	0.99	0.99	0.98	0.98	0.97	0.95	0.96	0.95

GAO: Genotype-Allelic-only; LAO: Local-Ancestry-only; APA: Ancestry Plus Allelic; LAAA: Local Ancestry Adjusted Allelic

TABLE 6 Median Pearson correlation coefficient between true and inferred local ancestry

Correlation	Genotype	African allele	African reference allele
All markers	1.00	0.81	0.97
Markers with overall MAF < 0.05	1.00	0.81	0.94
Markers with AFR MAF < 0.05	1.00	0.82	0.88

AFR: African reference population; MAF: Minor allele frequency.

ancestry effect = 0.50, and African-ancestry allelic effect = −1.04).

4 | DISCUSSION

Association studies in admixed populations bring opportunities to gain a deeper understanding of the genetic architecture of complex diseases and traits. At the same time, such studies require special treatment of the LA due to the differential origins of the chromosomal segments. In this study, we propose a robust and powerful two-step testing procedure, called the Local Ancestry Adjusted Allelic test or LAAA, for association analyses in admixed populations. We demonstrate its usefulness through extensive numerical simulations and real data analysis.

The underlying genetic architecture is unknown and can be complex in admixed populations. Therefore, we simulated

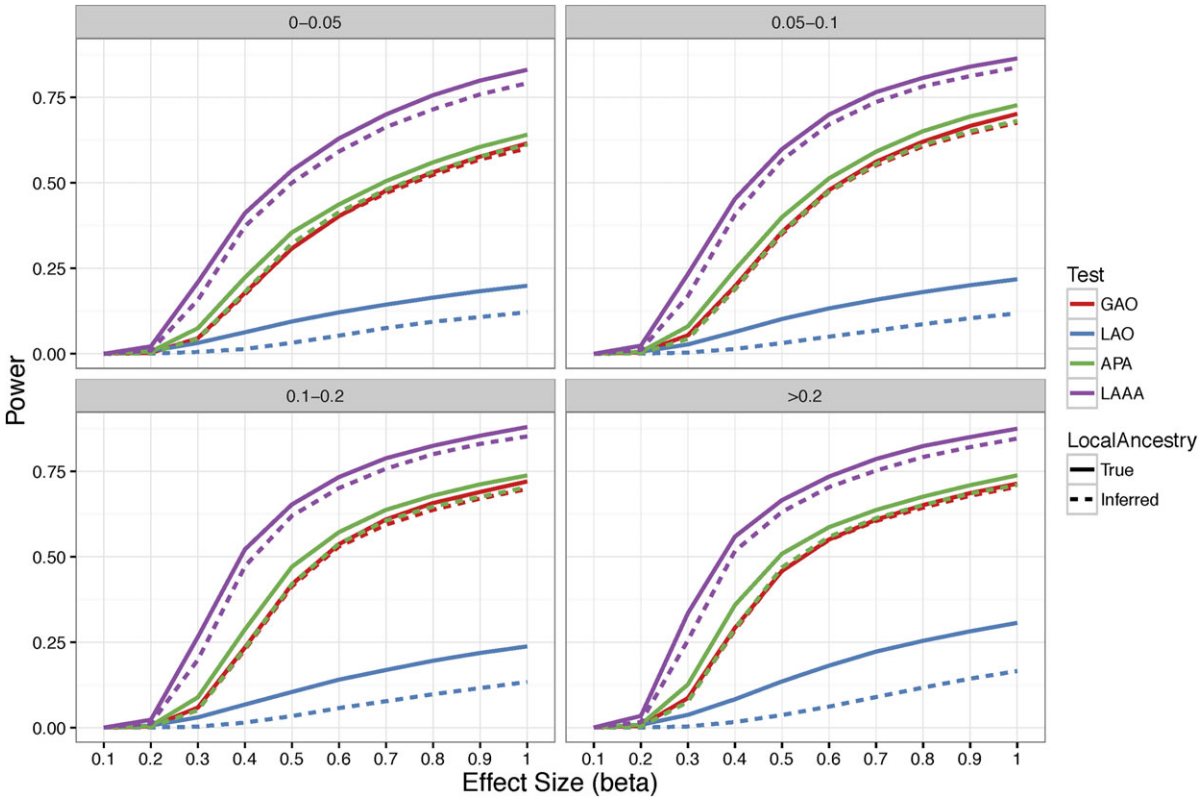

FIGURE 5 Power comparison with simulated true and inferred ancestry under scenario EH (EH: effect heterogeneity; GAO: Genotype-allelic-only; LAO: Local-ancestry-only; APA: Ancestry plus allelic; LAAA: Local ancestry adjusted allelic)

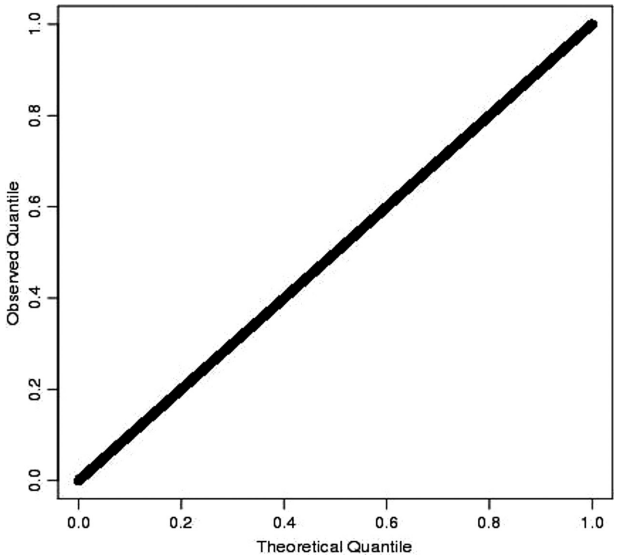

FIGURE 6 QQ plot of *P*-values obtained using LAAA under null

data covering four scenarios that reflect the complexity of real admixed data, including an ancestry only effect (scenario LA), both ancestry and allelic effects (scenario AA), a homogeneous allelic effect across ancestral populations (scenario GA) and heterogeneous allelic effects across ancestral populations (scenario EH). As shown in the results, without prior knowledge of the form of association, our proposed LAAA, which tests for allelic, ancestry, and the ancestry-specific allelic effects simultaneously, achieves competitive performance in all scenarios and particularly manifests its superiority when effect heterogeneity exists. As expected, power of all proposed tests depends on the true unknown underlying genetic architecture. Not surprisingly, power is the greatest when the testing procedure matches the underlying simulated model. For example, when an ancestry effect is the only source of association, the LAO test is more powerful test due to 1) its matching with the simulated/true model, and 2) the lower significant threshold (7×10^{-6}). As another example, when a homogeneous allelic effect is the only source of association, the GAO test has the best performance. Most importantly, though, LAAA is robust to the unknown underlying model: it either achieves the greatest power or is a close-match second across all scenarios.

LAAA is particularly powerful in the presence of effect heterogeneity. The power comparison (Tables 3 and 4) demonstrates the gain in power of the joint test in LAAA over BMIX and CPM. The enhanced power comes from our parameterization of effect heterogeneity using the number of ancestry-specific reference alleles, which is capturing the LA information down to the allele level. This advantage is more evident when LA is estimated, as the estimation accuracy of the number of ancestry-specific reference alleles is higher than that of the total number of ancestry-specific alleles (including the typically rarer alternative/nonreference allele), the latter of which is used in the cross-product term model (CPM).

We used a GWAS significance threshold of 5×10^{-8} for all tests involving allelic effects (GAO, APA, LAAA, and CPM) and 7×10^{-6} for LAO model, the test used in admixture mapping. The rationale for using a more lenient threshold for genome-wide significance of admixture mapping is based on previous theoretical analysis and simulation results that a threshold of 7×10^{-6} provides a genome-wide type I error of 0.05, because of the extensive regional correlation in LA in admixed populations (Pe'er, et al., 2008). However, it is still unclear whether we should adjust the GWAS significance threshold accordingly for other tests because of the reduced number of independent tests in admixed populations.

The LA and ancestry-specific alleles are estimated using HAPMIX and our assessment shows the estimation is highly accurate. However, the accuracy comes at the cost of high computational demand. To cope with the computational burden, we restricted our reference to the overlapping markers between the reference and genotyped markers, allowing only imputation for sporadic missingness among the genotypes at the typed markers. As a result, our analysis is limited to typed markers. To obtain a finer resolution, one may apply a suggestive *P*-value, e.g., 1×10^{-6}, to obtain the candidate region and follow it up by fine mapping. Through restricting to a small region, it is computationally feasible to infer the LA for

TABLE 7 Replication of published loci associated with hemoglobin using two-step testing procedure

SNP	Chr	Pos[1]	*P*-value	Allelic effect[2]	Ancestry effect[3]	African ancestry allele effect[4]	N[5]	Ref.
rs9940149	16	300641	7.23×10^{-7}	**1.87**	0.62	−1.24	5711	(Li, et al., 2013)
rs7199221	16	3101639	1.74×10^{-7}	**−3.89**	−0.73	2.59	5711	(Lo, et al., 2011)
rs1211375	16	240280	7.32×10^{-9}	0.86	−0.43	**−1.53**	5711	(Lo, et al., 2011)
Replication of the identified effect heterogeneity in an independent data set (WHI)								
rs1211375	16	240280	5.78×10^{-6}	0.54	0.50	**−1.04**	8087	(Lo, et al., 2011)

[1]SNP genomic position is shown in GRCh37.
[2]Test statistics associated with allelic effect.
[3]Test statistics associated with ancestry effect.
[4]Test statistics associated with allele specific allele effect.
[5]Sample size.

both typed and imputed markers. We learned from our analyses that it is necessary to raise the mutation rate parameter by a factor of 10 when using sequencing-based reference haplotypes from the 1000 genome project as is recommended in the HAPMIX tutorial. After using a higher mutation rate, we obtained similar LA estimates as reported previously (Reiner, et al., 2012).

We demonstrate the usefulness of our LAAA testing procedure through testing for the association with hemoglobin using data from CARe. In the first step, we replicated previous findings and in the second step, we were able to pinpoint the primary source of association. When a SNP exceeds genome-wide significance in the first step, it is advisable to check the MAF of the tested SNP in each of the reference populations (e.g., AFR and EUR in the 1000 Genomes project). We noticed that a marker with a MAF of zero in one reference population may be identified as significant with strong effect heterogeneity, which is actually an artifact due to LA estimation error. Our analysis replicated three previously reported SNPs on chromosome 16, rs9940149, rs7199221, and rs1211375. The strong allelic effects from rs9940149 and rs7199221 suggest similar effects across ancestral populations. The strong ancestry-specific allelic effect at rs1211375 suggests effect heterogeneity at this locus. Rs1211375, a common variant in both EUR and AFR (MAF = $\approx$ 0.3), has been reported to have a strong association with lower hemoglobin in African Americans but not in Caucasians. It has been suggested that this association might be driven by an African α-globin structure variation (CNVR6569), which is absent in Europeans (Lo, et al., 2011). This observation is consistent with the known copy number variations of α-globin genes that lead to α-thalassemia.

Our current two-step testing procedure is presented in admixed sample with two ancestral populations, such as African Americans. It may be extended to association studies in three-way admixed populations such as Hispanics/Latinos. However, several issues have to be addressed before such an adaptation. First, our method relies on the joint distribution of LA and allele status. However, such joint distribution is not readily available in the current LA inference tools for three-way admixed populations. Second, our method requires highly accurate LA estimation. Yet there is no consensus regarding the most accurate LA inference method with respect to three-way admixture. Therefore, the future direction of our current study includes method development for the inference of joint distribution of three-way admixed populations and extensive assessment of LA inference methods.

In summary, we developed a powerful and robust two-step testing procedure tailored for genetic association studies in admixed populations. Our method takes into account the presence of LA, thus achieving well-controlled type I error and enhanced power. We demonstrate the usefulness of our method in both simulation and real data analysis. We have implemented the proposed pipeline to facilitate LA adjusted association analysis, which is freely available online at: https://yunliweb.its.unc.edu/LAAA.

ACKNOWLEDGMENTS

The authors thank the WHI Study and the CARe project for generating the data. The research is supported by R01HG006292 (Y.L.), R01HG006703 (Y.L), R21HL126045 (L.R. and E.M.L.), and R01HL129132 (awarded to Y.L. and A.P.R).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Yun Li http://orcid.org/0000-0002-3467-2599

REFERENCES

Auer, P. L., Johnsen, J. M., Johnson, A. D., Logsdon, B. A., Lange, L. A., Nalls, M. A., … Li, Y. (2012). Imputation of exome sequence variants into population- based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project. *Am J Hum Genet*, *91*(5), 794–808.

Auer, P. L., & Teumer, A. (2014). Rare and low-frequency coding variants in CXCR2 and other genes are associated with hematological traits. *Nat Genet*, *46*, 629–634.

Baran, Y., Pasaniuc, B., Sankararaman, S., Torgerson, D. G., Gignoux, C., Eng, C., … Halperin, E. (2012). Fast and accurate inference of local ancestry in Latino populations. *Bioinformatics*, *28*, 1359–1367.

Bild, D. E., Bluemke, D. A., Burke, G. L., Detrano, R., Diez Roux, A. V., Folsom, A. R., … Tracy, R. P. (2002). Multi-ethnic study of atherosclerosis: Objectives and design. *Am J Epidemiol*, *156*, 871–881.

Chakraborty, R., & Weiss, K. M. (1988). Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc Natl Acad Sci U S A*, *85*(23), 9119–9123.

Chen, R., Corona, E., Sikora, M., Dudley, J. T., Morgan, A. A., Moreno-Estrada, A., … Butte, A. J. (2012). Type 2 diabetes risk alleles demonstrate extreme directional differentiation among human populations, compared to other diseases. *PLoS Genet*, *8*(4), e1002621.

Chen, Z., Tang, H., Qayyum, R., Schick, U. M., Nalls, M. A., Handsaker, R., … Reiner, A. P. (2013). Genome-wide association analysis of red blood cell traits in African Americans: The COGENT Network. *Hum Mol Genet*, *22*(12), 2529–2538.

Conneely, K. N., & Boehnke, M. (2007). So many correlated tests, so little time! Rapid adjustment of p values for multiple correlated tests. *American Journal of Human Genetics*, *81*, 1158–1168.

Frazier-Wood, A. C., Manichaikul, A., Aslibekyan, S., Borecki, I. B., Goff, D. C., Hopkins, P. N., … Arnett, D. K. (2013). Genetic variants associated with VLDL, *LDL and HDL particle size differ with race/ethnicity. Hum Genet*, *132*(4), 405–413.

Friedman, G. D., Cutter, G. R., Donahue, R. P., Hughes, G. H., Hulley, S. B., Jacobs, D. R., Jr., … Savage, P. J. (1988). CARDIA: Study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*, *41*(11), 1105–1116.

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., … Pollak, M. R. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science*, *329*(5993), 841–845.

Helgason, A., Palsson, S., Thorleifsson, G., Grant, S. F., Emilsson, V., Gunnarsdottir, S., … Stefansson, K. (2007). Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet*, *39*(2), 218–225.

International HapMap, C. (2005). A haplotype map of the human genome. *Nature*, *437*, 1299–1320.

Ioannidis, J. P., Ntzani, E. E., & Trikalinos, T. A. (2004). 'Racial' differences in genetic effects for complex diseases. *Nature Genetics*, *36*, 1312–1318.

Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., … Kamatani, N. (2010). Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat Genet*, *42*(3), 210–215.

Kittles, R. A., Chen, W., Panguluri, R. K., Ahaghotu, C., Jackson, A., Adebamowo, C. A., … Dunston, G. M. (2002). CYP3A4-V and prostate cancer in African Americans: Causal or confounding association because of population stratification? *Hum Genet*, *110*(6), 553–560.

Lamason, R. L., Mohideen, M. A., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C., … Cheng, K. C. (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, *310*(5755), 1782–1786.

Levin, A. M., Iannuzzi, M. C., Montgomery, C. G., Trudeau, S., Datta, I., Adrianto, I., … Rybicki, B. A. (2014). Admixture fine-mapping in African Americans implicates XAF1 as a possible sarcoidosis risk gene. *PLoS One*, *9*(3), e92646.

Li, J., Glessner, J. T., Zhang, H., Hou, C., Wei, Z., Bradfield, J. P., … Sleiman, P. M. (2013). GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. *Hum Mol Genet*, *22*(7), 1457–1464.

Liu, J., Lewinger, J. P., Gilliland, F. D., Gauderman, W. J., & Conti, D. V. (2013). Confounding and heterogeneity in genetic association studies with admixed populations. *Am J Epidemiol*, *177*(4), 351–360.

Lo, K. S., Wilson, J. G., Lange, L. A., Folsom, A. R., Galarneau, G., Ganesh, S. K., … Lettre, G. (2011). Genetic association analysis highlights new loci that modulate hematological trait variation in Caucasians and African Americans. *Hum Genet*, *129*(3), 307–317.

Loth, D. W., Artigas, M. S., Gharib, S. A., Wain, L. V., Franceschini, N., Koch, B., … London, S. J. (2014). Genome-wide association analysis identifies six new loci associated with forced vital capacity. *Nat Genet*, *46*(7), 669–677.

Mao, X., Li, Y., Liu, Y., Lange, L., & Li, M. (2013). Testing genetic association with rare variants in admixed populations. *Genet Epidemiol*, *37*(1), 38–47.

McHugh, C., Brown, L., & Thornton, T. A. (2016). Detecting Heterogeneity in Population Structure Across the Genome in Admixed Populations. *Genetics*, *204*(1), 43–56.

Musunuru, K., Lettre, G., Young, T., Farlow, D. N., Pirruccello, J. P., Ejebe, K. G., … Resource, N. C. G.A. (2010). Candidate gene association resource (CARe): Design, methods, and proof of concept. *Circ Cardiovasc Genet*, *3*(3), 267–275.

Nalls, M. A., Wilson, J. G., Patterson, N. J., Tandon, A., Zmuda, J. M., Huntsman, S., … Ziv, E. (2008). Admixture mapping of white cell count: Genetic locus responsible for lower white blood cell count in the Health ABC and Jackson Heart studies. *Am J Hum Genet*, *82*(1), 81–87.

Parra, E. J., Marcini, A., Akey, J., Martinson, J., Batzer, M. A., Cooper, R., … Shriver, M. D. (1998). Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet*, *63*(6), 1839–1851.

Pasaniuc, B., Zaitlen, N., Lettre, G., Chen, G. K., Tandon, A., Kao, W. H., … Price, A. L. (2011). Enhanced statistical tests for GWAS in admixed populations: Assessment using African Americans from CARe and a Breast Cancer Consortium. *PLoS Genet*, *7*(4), e1001371.

Patterson, N., Hattangadi, N., Lane, B., Lohmueller, K. E., Hafler, D. A., Oksenberg, J. R., … Reich, D. (2004). Methods for high-density admixture mapping of disease genes. *Am J Hum Genet*, *74*(5), 979–1000.

Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol*, *32*(4), 381–385.

Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature*, *538*, 161–164.

Price, A. L., Tandon, A., Patterson, N., Barnes, K. C., Rafaels, N., Ruczinski, I., … Myers, S. (2009). Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genetics*, *5*(6), e1000519.

Qin, H., Morris, N., Kang, S. J., Li, M., Tayo, B., Lyon, H., … Zhu, X. (2010). Interrogating local population structure for fine mapping in genome-wide association studies. *Bioinformatics*, *26*(23), 2961–2968.

Reiner, A. P., Beleza, S., Franceschini, N., Auer, P. L., Robinson, J. G., Kooperberg, C., … Tang, H. (2012). Genome-wide association and population genetic analysis of C-reactive protein in African American and Hispanic American women. *Am J Hum Genet*, *91*(3), 502–512.

Reiner, A. P., Lettre, G., Nalls, M. A., Ganesh, S. K., Mathias, R., Austin, M. A., … Wilson, J. G. (2011). Genome-wide association study of white blood cell count in 16,388 African Americans: The continental origins and genetic epidemiology network (COGENT). *PLoS Genet*, *7*(6), e1002108.

Rosenberg, N. A., Huang, L., Jewett, E. M., Szpiech, Z. A., Jankovic, I., & Boehnke, M. (2010). Genome-wide association studies in diverse populations. *Nature Reviews Genetics*, *11*(5), 356–366.

Schaffner, S. F., Foo, C., Gabriel, S., Reich, D., Daly, M. J., & Altshuler, D. (2005). Calibrating a coalescent simulation of human genome sequence variation. *Genome Res*, *15*(11), 1576–1583.

Shriner, D. (2013). Overview of admixture mapping. *Current protocols in human genetics / editorial board, Jonathan L. Haines … [et al.]*, Chapter 1, Unit 1 23.

Shriner, D., Adeyemo, A., & Rotimi, C. N. (2011). Joint ancestry and association testing in admixed individuals. *PLoS Computational Biology*, *7*, e1002325.

Tang, H., Siegmund, D. O., Johnson, N. A., Romieu, I., & London, S. J. (2010). Joint testing of genotype and ancestry association in admixed families. *Genet Epidemiol*, *34*(8), 783–791.

Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., … Kathiresan, S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, *466*(7307), 707–713.

Thornton, T. A., & Bermejo, J. L. (2014). Local and global ancestry inference and applications to genetic association analysis for admixed populations. *Genetic Epidemiology*, *38*(Suppl 1), S5–S12.

van der Harst, P., Zhang, W., Mateo Leach, I., Rendon, A., Verweij, N., Sehmi, J., … Chambers, J. C. (2012). Seventy-five genetic loci influencing the human red blood cell. *Nature*, *492*(7429), 369–375.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *Am J Hum Genet*, *90*(1), 7–24.

Wang, C. L., Zhan, X. W., Bragg-Gresham, J., Kang, H. M., Stambolian, D., Chew, E. Y., … Study, F. (2014). Ancestry estimation and control of population stratification for sequence-based association studies. *Nature Genetics*, *46*(4), 409.

Wang, X., Zhu, X., Qin, H., Cooper, R. S., Ewens, W. J., Li, C., & Li, M. (2011). Adjustment for local ancestry in genetic association analysis of admixed populations. *Bioinformatics*, *27*(5), 670–677.

Wegmann, D., Kessner, D. E., Veeramah, K. R., Mathias, R. A., Nicolae, D. L., Yanek, L. R., … Novembre, J. (2011). Recombination rates in admixed individuals identified by ancestry-based inference. *Nat Genet*, *43*(9), 847–853.

Winkler, C. A., Nelson, G. W., & Smith, M. W. (2010). Admixture mapping comes of age. *Annual review of genomics and human genetics*, *11*, 65–89.

Zhang, J., & Stram, D. O. (2014). The role of local ancestry adjustment in association studies using admixed populations. *Genetic Epidemiology*, *38*, 502–515.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Duan Q, Xu Z, Raffield L, et al. A robust and powerful two-step testing procedure for local ancestry adjusted allelic association analysis in admixed populations. *Genet Epidemiol*. 2018;42:288–302. https://doi.org/10.1002/gepi.22104